

SEQUENCE LISTING

<110> Japan Science and Technology Corporation

<120> Human nucleoprotein having a WW domain and
a polynucleotide encoding the protein

<130> 00-F-061PCT

<140> PCT/JP00/08253

<141> 2000-11-22

<150> JP11-332572

<151> 1999-11-24

<160> 7

<170> PatentIn Ver. 2.0

<210> 1

<211> 704

<212> PRT

<213> Homo sapiens

<400> 1

Met Ala Asn Glu Asn His Gly Ser Pro Arg Glu Glu Ala Ser Leu Leu

1 5 10 15

Ser His Ser Pro Gly Thr Ser Asn Gln Ser Gln Pro Cys Ser Pro Lys

20 25 30

Pro Ile Arg Leu Val Gln Asp Leu Pro Glu Glu Leu Val His Ala Gly

09889722 101504

[illegible]

Ile Arg Leu Ser Arg Ile Lys Phe Arg Glu Glu Ala Lys Arg Leu Leu
 275 280 285
 Phe Lys Tyr Ala Glu Ala Ala Arg Arg Leu Ile Glu Ser Arg Ser Ala
 290 295 300
 Ser Pro Asp Ser Arg Lys Val Val Lys Trp Asn Val Glu Asp Thr Phe
 305 310 315 320
 Ser Trp Leu Arg Lys Asp His Ser Ala Ser Lys Glu Asp Tyr Met Asp
 325 330 335
 Arg Leu Glu His Leu Arg Arg Gln Cys Gly Pro His Val Ser Ala Ala
 340 345 350
 Ala Lys Asp Ser Val Glu Gly Ile Cys Ser Lys Ile Tyr His Ile Ser
 355 360 365
 Leu Glu Tyr Val Lys Arg Ile Arg Glu Lys His Leu Ala Ile Leu Lys
 370 375 380
 Glu Asn Asn Ile Ser Glu Glu Val Glu Ala Pro Glu Val Glu Pro Arg
 385 390 395 400
 Leu Val Tyr Cys Tyr Pro Val Arg Leu Ala Val Ser Ala Pro Pro Met
 405 410 415
 Pro Ser Val Glu Met His Met Glu Asn Asn Val Val Cys Ile Arg Tyr
 420 425 430
 Lys Gly Glu Met Val Lys Val Ser Arg Asn Tyr Phe Ser Lys Leu Trp
 435 440 445
 Leu Leu Tyr Arg Tyr Ser Cys Ile Asp Asp Ser Ala Phe Glu Arg Phe
 450 455 460
 Leu Pro Arg Val Trp Cys Leu Leu Arg Arg Tyr Gln Met Met Phe Gly
 465 470 475 480
 Val Gly Leu Tyr Glu Gly Thr Gly Leu Gln Gly Ser Leu Pro Val His
 485 490 495
 Val Phe Glu Ala Leu His Arg Leu Phe Gly Val Ser Phe Glu Cys Phe

0989722101901
 T06TOT2276960

$\langle 211 \rangle$ 2112

<212> DNA

<213> Homo sapiens

<400> 2

atggccaatg agaatcacgg cagcccccgaggaggaagcgt ccctgctgag tcaactcccca 60
 ggtacctcca atcagagcca gccctgttct ccaaagccaa tccgcctggt tcaggacctc 120
 ccagaggagc tggatcatgc aggcctgggag aagtgtgga gccggaggga gaatcgtccc 180
 tactacttca accgattcac caaccagtcc ctgtgggaga tggcctgct ggggcagcac 240
 gatgtgattt cggacccttt ggggctgaat gcgacccac tggcccaaga ctcaagcttg 300
 gtggaaactc ccccggtga gaacaagccc agaaagcggc agctctcgga agagcagcca 360
 agcggcaatg gtgtgaagaa gccaagatt gaaatcccag tgacaccac aggccagtcg 420
 gtgcccagct ccccagtat ccaggaacc ccaacgtga agatgtggg tacgtccct 480
 gaagataaac agcaggcagc tctctacga cccactgagg tctactgga cctggacatc 540
 cagaccaatg ctgtcatcaa gcaccgggg ccttcagagg tgctgcccc gcacccgaa 600
 gtggaactgc tccgtctca gctatcctg aagcttcggc agcactatcg ggagctgtgc 660
 cagcagcgag agggcattga gcctccacgg gactcttca accgctggat gctggagcgc 720
 aaggtggtag acaaaggatc tgacccctg ttgccagca actgtgaacc agtcgtgtca 780
 ccttccatgt ttctgaaat catgaacgac attcctatca ggttatccg aatcaagttc 840
 cgggaggaag ccaagcgctt gctctttaa tatgcggagg ccgccaggcg gctcatcgag 900
 tccaggagtg catccctga cagtaggaag gtggtcaa atggaatgtga agacacctt 960
 agctggcttc ggaaggacca ctacgctcc aaggaggact acatggatcg cctggagcat 1020
 ctgcggaggc agtgtggccc ccacgtctcg gccgcagcca aggactcgt ggaaggcatc 1080
 tgcagtaaga tctaccacat ctccctggag tacgtcaaac ggatccgaga gaagcacctt 1140
 gccatcctca aggaaaacaa catctcagag gaggtggagg cccctgaggt ggagccccgc 1200
 ctagtgtact gctaccagc cggctggct gtgtctgcac cgccatgcc cagcgtggag 1260
 atgcacatgg agaacaacgt ggtctgcatc cgggtataagg gagagatggt caaggtcagc 1320
 cgcaactact tcagcaagct gtggctcctt taccgtaca gctgcattga tgactctgcc 1380
 tttagaggtt tcctgcccc ggtctggtgt cttctccgac ggtaccagat gatgttcggc 1440
 gtggcctct acgaggggac tggcctgcag ggatcgctgc ctgtgcatgt ctttgaggcc 1500

0989723 101901

ctccaccgac tctttggcgt cagcttcgag tgccttcgcct caccctcaa ctgctacttc 1560
 cgccagtact gttctgcctt ccccgacaca gacggctact ttggctcccg cgggccctgc 1620
 ctagactttg ctccactgag tggttcattt gaggccaacc ctcccttctg cgaggagctc 1680
 atggatgcca tggctctctca ctttgagaga ctgcttgaga gctcaccgga gccctgtcc 1740
 ttcatcgtgt tcatccctga gtggcgggaa ccccaaacac cagcgctcac ccgcatggag 1800
 cagagccgct tcaaacgcca ccagttgatc ctgcctgcct ttgagcatga gtaccgcagt 1860
 ggctcccagc acatctgcaa gaaggaggaa atgcactaca aggccgtcca caacacggct 1920
 gtgctcttcc tacagaacga ccttggcttt gccaaagtgg cgccgacgcc tgaacggctg 1980
 caggagctga gtgctgccta ccggcagtc ggcgcagcc acagctctgg ttcttcctca 2040
 tcgtctcct cggaggccaa ggaccgggac tcgggccgtg agcaggtcc tagccgcgag 2100
 cctcacccca ct 2112

<210> 3

<211> 2669

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (292).. (2406)

<400> 3

acacaagatg gcggcagcgg cgctggggag ggcgaggcgg aggcggcaaa acgggcggtc 60
 gagcagaacg ttagccgcg tccctccag tccgtccgg gcagctgctg atgcaaggaa 120
 tccctgggc tccgtccac tccactgctg accagcccat tcgctgtgc tgagtcttcc 180
 tgcaggcctt tccttgctc tgtgggacc tgtgggggtc catccggctg gagaagaaaa 240
 gcctctcatg ctaacgttgc agacccaga gggtcctgtg tgggtgtgga g atg gcc 297

Met Ala

0909722-101904

1

aat gag aat cac ggc agc ccc cgg gag gaa gcg tcc ctg ctg agt cac 345
 Asn Glu Asn His Gly Ser Pro Arg Glu Glu Ala Ser Leu Leu Ser His
 5 10 15
 tcc cca ggt acc tcc aat cag agc cag ccc tgt tct cca aag cca atc 393
 Ser Pro Gly Thr Ser Asn Gln Ser Gln Pro Cys Ser Pro Lys Pro Ile
 20 25 30
 cgc ctg gtt cag gac ctc cca gag gag ctg gtg cat gca ggc tgg gag 441
 Arg Leu Val Gln Asp Leu Pro Glu Glu Leu Val His Ala Gly Trp Glu
 35 40 45 50
 aag tgc tgg agc cgg agg gag aat cgt ccc tac tac ttc aac cga ttc 489
 Lys Cys Trp Ser Arg Arg Glu Asn Arg Pro Tyr Tyr Phe Asn Arg Phe
 55 60 65
 acc aac cag tcc ctg tgg gag atg ccc gtg ctg ggg cag cac gat gtg 537
 Thr Asn Gln Ser Leu Trp Glu Met Pro Val Leu Gly Gln His Asp Val
 70 75 80
 att tcg gac cct ttg ggg ctg aat gcg acc cca ctg ccc caa gac tca 585
 Ile Ser Asp Pro Leu Gly Leu Asn Ala Thr Pro Leu Pro Gln Asp Ser
 85 90 95
 agc ttg gtg gaa act ccc ccg gct gag aac aag ccc aga aag cgg cag 633
 Ser Leu Val Glu Thr Pro Pro Ala Glu Asn Lys Pro Arg Lys Arg Gln
 100 105 110
 ctc tcg gaa gag cag cca agc ggc aat ggt gtg aag aag ccc aag att 681
 Leu Ser Glu Glu Gln Pro Ser Gly Asn Gly Val Lys Lys Pro Lys Ile
 115 120 125 130
 gaa atc cca gtg aca ccc aca ggc cag tcg gtg ccc agc tcc ccc agt 729
 Glu Ile Pro Val Thr Pro Thr Gly Gln Ser Val Pro Ser Ser Pro Ser
 135 140 145
 atc cca gga acc cca acg ctg aag atg tgg ggt acg tcc cct gaa gat 777

09669762-104904
 105404-22459350

Ile Pro Gly Thr Pro Thr Leu Lys Met Trp Gly Thr Ser Pro Glu Asp
 150 155 160
 aaa cag cag gca gct ctc cta cga ccc act gag gtc tac tgg gac ctg 825
 Lys Gln Gln Ala Ala Leu Leu Arg Pro Thr Glu Val Tyr Trp Asp Leu
 165 170 175
 gag atc cag acc aat gct gtc atc aag cac cgg ggg cct tca gag gtg 873
 Asp Ile Gln Thr Asn Ala Val Ile Lys His Arg Gly Pro Ser Glu Val
 180 185 190
 ctg ccc ccg cat ccc gaa gtg gaa ctg ctc cgc tct cag ctc atc ctg 921
 Leu Pro Pro His Pro Glu Val Glu Leu Leu Arg Ser Gln Leu Ile Leu
 195 200 205 210
 aag ctt cgg cag cac tat cgg gag ctg tgc cag cag cga gag ggc att 969
 Lys Leu Arg Gln His Tyr Arg Glu Leu Cys Gln Gln Arg Glu Gly Ile
 215 220 225
 gag cct cca cgg gag tct ttc aac cgc tgg atg ctg gag cgc aag gtg 1017
 Glu Pro Pro Arg Glu Ser Phe Asn Arg Trp Met Leu Glu Arg Lys Val
 230 235 240
 gta gac aaa gga tct gac ccc ctg ttg ccc agc aac tgt gaa cca gtc 1065
 Val Asp Lys Gly Ser Asp Pro Leu Leu Pro Ser Asn Cys Glu Pro Val
 245 250 255
 gtg tca cct tcc atg ttt cgt gaa atc atg aac gac att cct atc agg 1113
 Val Ser Pro Ser Met Phe Arg Glu Ile Met Asn Asp Ile Pro Ile Arg
 260 265 270
 tta tcc cga atc aag ttc cgg gag gaa gcc aag cgc ctg ctc ttt aaa 1161
 Leu Ser Arg Ile Lys Phe Arg Glu Glu Ala Lys Arg Leu Leu Phe Lys
 275 280 285 290
 tat gcg gag gcc gcc agg cgg ctc atc gag tcc agg agt gca tcc cct 1209
 Tyr Ala Glu Ala Ala Arg Arg Leu Ile Glu Ser Arg Ser Ala Ser Pro
 295 300 305

0000722.101504
 106107 22259900

gac agt agg aag gtg gtc aaa tgg aat gtg gaa gac acc ttt agc tgg	1257
Asp Ser Arg Lys Val Val Lys Trp Asn Val Glu Asp Thr Phe Ser Trp	
310 315 320	
ctt cgg aag gac cac tca gcc tcc aag gag gac tac atg gat cgc ctg	1305
Leu Arg Lys Asp His Ser Ala Ser Lys Glu Asp Tyr Met Asp Arg Leu	
325 330 335	
gag cat ctg cgg agg cag tgt ggc ccc cac gtc tcg gcc gca gcc aag	1353
Glu His Leu Arg Arg Gln Cys Gly Pro His Val Ser Ala Ala Ala Lys	
340 345 350	
gac tcc gtg gaa ggc atc tgc agt aag atc tac cac atc tcc ctg gag	1401
Asp Ser Val Glu Gly Ile Cys Ser Lys Ile Tyr His Ile Ser Leu Glu	
355 360 365 370	
tac gtc aaa cgg atc cga gag aag cac ctt gcc atc ctc aag gaa aac	1449
Tyr Val Lys Arg Ile Arg Glu Lys His Leu Ala Ile Leu Lys Glu Asn	
375 380 385	
aac atc tca gag gag gtg gag gcc cct gag gtg gag ccc cgc cta gtg	1497
Asn Ile Ser Glu Glu Val Glu Ala Pro Glu Val Glu Pro Arg Leu Val	
390 395 400	
tac tgc tac cca gtc cgg ctg gct gtg tct gca ccg ccc atg ccc agc	1545
Tyr Cys Tyr Pro Val Arg Leu Ala Val Ser Ala Pro Pro Met Pro Ser	
405 410 415	
gtg gag atg cac atg gag aac aac gtg gtc tgc atc cgg tat aag gga	1593
Val Glu Met His Met Glu Asn Asn Val Val Cys Ile Arg Tyr Lys Gly	
420 425 430	
gag atg gtc aag gtc agc cgc aac tac ttc agc aag ctg tgg ctc ctt	1641
Glu Met Val Lys Val Ser Arg Asn Tyr Phe Ser Lys Leu Trp Leu Leu	
435 440 445 450	
tac cgc tac agc tgc att gat gac tct gcc ttt gag agg ttc ctg ccc	1689
Tyr Arg Tyr Ser Cys Ile Asp Asp Ser Ala Phe Glu Arg Phe Leu Pro	

455	460	465	
cgg gtc tgg tgt ctt ctc cga cgg tac cag atg atg ttc ggc gtg ggc			1737
Arg Val Trp Cys Leu Leu Arg Arg Tyr Gln Met Met Phe Gly Val Gly			
470	475	480	
ctc tac gag ggg act ggc ctg cag gga tcg ctg cct gtg cat gtc ttt			1785
Leu Tyr Glu Gly Thr Gly Leu Gln Gly Ser Leu Pro Val His Val Phe			
485	490	495	
gag gcc ctc cac cga ctc ttt ggc gtc agc ttc gag tgc ttc gcc tca			1833
Glu Ala Leu His Arg Leu Phe Gly Val Ser Phe Glu Cys Phe Ala Ser			
500	505	510	
ccc ctc aac tgc tac ttc cgc cag tac tgt tct gcc ttc ccc gac aca			1881
Pro Leu Asn Cys Tyr Phe Arg Gln Tyr Cys Ser Ala Phe Pro Asp Thr			
515	520	525	530
gac ggc tac ttt ggc tcc cgc ggg ccc tgc cta gac ttt gct cca ctg			1929
Asp Gly Tyr Phe Gly Ser Arg Gly Pro Cys Leu Asp Phe Ala Pro Leu			
535	540	545	
agt ggt tca ttt gag gcc aac cct ccc ttc tgc gag gag ctc atg gat			1977
Ser Gly Ser Phe Glu Ala Asn Pro Pro Phe Cys Glu Glu Leu Met Asp			
550	555	560	
gcc atg gtc tct cac ttt gag aga ctg ctt gag agc tca ccg gag ccc			2025
Ala Met Val Ser His Phe Glu Arg Leu Leu Glu Ser Ser Pro Glu Pro			
565	570	575	
ctg tcc ttc atc gtg ttc atc cct gag tgg cgg gaa ccc cca aca cca			2073
Leu Ser Phe Ile Val Phe Ile Pro Glu Trp Arg Glu Pro Pro Thr Pro			
580	585	590	
gcg ctc acc cgc atg gag cag agc cgc ttc aaa cgc cac cag ttg atc			2121
Ala Leu Thr Arg Met Glu Gln Ser Arg Phe Lys Arg His Gln Leu Ile			
595	600	605	610
ctg cct gcc ttt gag cat gag tac cgc agt ggc tcc cag cac atc tgc			2169

09039722104904

Leu Pro Ala Phe Glu His Glu Tyr Arg Ser Gly Ser Gln His Ile Cys
 615 620 625
 aag aag gag gaa atg cac tac aag gcc gtc cac aac acg gct gtg ctc 2217
 Lys Lys Glu Glu Met His Tyr Lys Ala Val His Asn Thr Ala Val Leu
 630 635 640
 ttc cta cag aac gac cct ggc ttt gcc aag tgg gcg ccg acg cct gaa 2265
 Phe Leu Gln Asn Asp Pro Gly Phe Ala Lys Trp Ala Pro Thr Pro Glu
 645 650 655
 cgg ctg cag gag ctg agt gct gcc tac cgg cag tca ggc cgc agc cac 2313
 Arg Leu Gln Glu Leu Ser Ala Ala Tyr Arg Gln Ser Gly Arg Ser His
 660 665 670
 agc tct ggt tct tcc tca tcg tcc tcc tcg gag gcc aag gac cgg gac 2361
 Ser Ser Gly Ser Ser Ser Ser Ser Ser Ser Glu Ala Lys Asp Arg Asp
 675 680 685 690
 tcg ggc cgt gag cag ggt cct agc cgc gag cct cac ccc act taa 2406
 Ser Gly Arg Glu Gln Gly Pro Ser Arg Glu Pro His Pro Thr
 695 700 705
 catatcctgc ggggaggagg agccccaggg gtgctagtct ggactgctgg gactcgggcc 2466
 cctggggcct cagagggacc ccggctgcca ctgacatatg aagattatgg ttctgccagg 2526
 gctcccctcc ctgcctgtcc ccaagtcctc acctcaact ccctccaagt cccatgtata 2586
 taggtcctga tgccttccca accccgcccc tcaccctgtt gccaccttgt ttcatttgta 2646
 aaaggaaata cagaaacccc ccc 2669

<210> 4

<211> 26

<212> DNA

<213> Artificial sequence

09889722-401904

<220>

<213> Synthesized oligonucleotide

<400> 4

ccgaattcat ggccaatgag aatcac

26

<210> 5

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<213> Synthesized oligonucleotide

<400> 5

ccgtcgactt aagtgggggtg aggctc

26

<210> 6

<211> 33

<212> DNA

<213> Artificial sequence

<220>

<213> Synthesized oligonucleotide

<400> 6

cgaggatccg ttcaggacct cccagaggacg cta

33

090921 101001

<210> 7

<211> 33

<212> DNA

<213> Artificial sequence

<220>

<213> Synthesized oligonucleotide

<400> 7

cgagaattcc gaaatcacat cgtgctgccc cag

33

00889722-101904